

Post-print version: Molinuevo, B., Martínez-Membrives, E., Pera-Guardiola, V., Requena, A., Torrent, N., Bonillo, A., Batalla, I., Torrubia, R., & Frick, P. J. (2020). Psychometric Properties of the Clinical Assessment of Prosocial Emotions: Version 1.1 (CAPE 1.1) in Young Males Who Were Incarcerated. *Criminal Justice and Behavior*, 47 (5), 547–563. doi: 10.1177/0093854819892931

**Psychometric Properties of the Clinical Assessment of Prosocial Emotions: Version 1.1
(CAPE 1.1) in Young Males Who Were Incarcerated**

Beatriz Molinuevo^{a,b}, Esther Martínez-Membrives^{a,b}, Vanessa Pera-Guardiola^{c,d,e}, Albert Requena^f, Noemí Torrent^g, Albert Bonillo^h, Iolanda Batalla^{d,e,i}, Rafael Torrubia^{a,b} and Paul J. Frick^{j,k}

^a Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona
(UAB), Barcelona, Spain

^b Institute of Neurosciences (INc), UAB, Barcelona, Spain

^c Child and Adolescent Mental Health Service (CAMHS), Sant Joan de Deu Terres de Lleida,
Spain

^d Biomedical Research Institute (IRB), Lleida, Spain

^e Department of Medicine, Universitat de Lleida, Lleida, Spain

^f Unit of Attention and Evaluation to Affected by Terrorism, Barcelona, Spain

^g PsicoGarrigues, Lleida, Spain

^h Department of Psychobiology and Methodology of Health Sciences
UAB, Barcelona, Spain

ⁱ Santa Maria University Hospital, Department of Psychiatry, Lleida, Spain

^j Department of Psychology, Louisiana State University, Louisiana, USA

^k Learning Science Institute of Australia, Brisbane, Australia

Acknowledgements. The authors are grateful to all the participating youths and juvenile justice professionals of the detention centers L'Alzina (Barcelona) and El Segre (Lleida). They would like to recognize Xènia Blaya, BS, Juanjo J. Martínez, BS, David Garreta, MS, Anna Camarasa, BS, Carmen Tello, PhD, and Manel Fortes, BS, for assistance in data collection. This research was financially supported by the *Centre d'Estudis Jurídics i Formació Especialitzada, Generalitat de Catalunya* (DOGC Núm. 7024-23.12.05; DOGC DOGC Núm. 7274 – 27/12/16), by the *Departament d'Economia i Coneixement, Generalitat de Catalunya* (2014SGR-1587), and by the *Ministerio de Economía, Industria y Competitividad*, Spanish Government (PSI2015-67441-R). The funding organizations had no further role in the study design; collection, analysis, and interpretation of data; writing of the report; and decision to submit the manuscript for publication.

The authors declare that Dr. Paul Frick is the creator of the CAPE 1.1 and they declare no other conflicts of interest.

Author's note

Correspondence concerning this article should be addressed to B. Molinuevo,
Departament de Psiquiatria i Medicina Legal. Facultat de Medicina. Universitat Autònoma de
Barcelona. Avinguda de Can Domènech, s/n, 08193 Bellaterra (Cerdanyola del Vallès).
Barcelona, Spain. Phone number: +34 93 581 1223. Fax: +34 93 581 1435. E-mail:
beatriz.molinuevo@uab.cat

Biographical sketches

Beatriz Molinuevo is professor in the Department of Psychiatry and Forensic Medicine at the Autonomous University of Barcelona. Her main research interests focuses on the emotional and cognitive correlates of psychopathic traits in children and adolescents and on studying the influence of parenting practices in socialization processes.

Esther Martínez-Membrives is school counselor and teacher of Vocational Training at the Department of Education, Government of Catalonia. Her research interests include fear and anxiety and emotional learning and psychopathy in youth population.

Vanessa Pera-Guardiola is child psychiatrist; the assistance director of Child and Youth Mental Health Service at Sant Joan de Déu Terres de Lleida and assistant professor at the Department of Psychology at Lleida University. She is a research member of Biomedical Research Institute (IRB), Lleida. Her research focus on psychopathy and conduct disorders in youth population.

Albert Requena is forensic psychologist and clinical child psychologist, specialized in child abuse and children with autism spectrum disorder. He is a trauma psychologist at the Unit of Attention and Evaluation to Affected by Terrorism in Barcelona.

Noemí Torrent is child psychologist at PsicoGarrigues. Her research interests include the psychopathy and behavioral disorders in child and youth population.

Albert Bonillo is professor of statistical data analysis at the Department of Psychobiology and Methodology of Health Sciences at the Autonomous University of Barcelona. His research

and methodological works focus on data-management and data-cleaning in scientific studies and his applied research focus on executive function and the measure of psychopathy.

Iolanda Batalla is psychiatrist in Santa Maria University Hospital, Associate Professor of Medicine at the Department of Medicine at the University of Lleida and researcher at the Biomedical Research Institute of Lleida. Her clinical practice and research focuses on personality disorders, and include the study and treatment of psychopathy and borderline personality disorders, from childhood to adult age.

Rafael Torrubia is professor in the Department of Psychiatry and Forensic Medicine at the Autonomous of Barcelona. His research focuses on the study of temperamental traits related with vulnerability to externalizing disorders.

Paul J. Frick is the Roy Crumpler Memorial Chair and Professor in the Department of Psychology at Louisiana State University and Professorial Fellow in the Institute for Learning Sciences and Teacher Education at Australian Catholic University. His research focuses on advancing knowledge on the developmental pathways that can lead to serious antisocial and aggressive behavior and on using this research to enhance the assessment and treatment of youth with serious behavior problems.

Abstract

The Clinical Assessment of Prosocial Emotions: Version 1.1 (CAPE 1.1) uses structured clinical judgements to diagnose the “with Limited Prosocial Emotions” specifier for Conduct Disorder. This study examined: (a) the internal consistency and inter-rater agreement and (b) the convergent and divergent validity of the CAPE 1.1 in 72 young males who were incarcerated in two Spanish juvenile detention centers (age range = 14-22 years). The CAPE 1.1 showed good inter-rater agreement for making the diagnosis of the specifier and adequate internal consistency. The CAPE 1.1 was associated with other measures of callous-unemotional traits but less consistently associated with other dimensions of psychopathy. Youth who met diagnostic criteria for the specifier scored higher on externalizing problems but did not differ from other youth who were incarcerated on internalizing problems. These results provide preliminary support for the psychometric properties of the CAPE 1.1 for the clinical assessment of the specifier.

Keywords: Clinical Assessment of Prosocial Emotions (CAPE 1.1), callous-unemotional traits, limited prosocial emotions, juvenile delinquency.

**Psychometric Properties of the Clinical Assessment of Prosocial Emotions (CAPE 1.1)
in Young Males Who Were Incarcerated**

Callous-unemotional (CU) traits define the affective components of psychopathy in adult samples (Hare & Neumann, 2008) and the affective components of conscience in child samples (Frick, Ray, Thornton, & Kahn, 2014a). Further, there is now substantial evidence to support the importance of CU traits for designating a clinically important subgroup of antisocial youth (Frick, Stickle, Dandreaux, Farrell, & Kimonis; 2005). That is, recent reviews of the available literature have shown that the presence of elevated CU traits in children and adolescents with serious behavior problems designates a group that is especially severe, violent, and difficult to treat (Blair, Leibenluft, & Pine, 2014; Frick, Ray, Thornton, & Kahn, 2014b; Herpers, Rommelse, Bons, Buitelaar, & Scheepers, 2012). These reviews have also indicated that children and adolescents with elevated CU traits show a number of distinct genetic, biological, emotional, cognitive, and social characteristics when compared to antisocial youth who are not elevated on these traits, suggesting that the causal processes underlying the behavior problems of these two groups may be different (Frick et al., 2014b).

Based on this research, the most recent edition of *The Diagnostic and Statistical Manual of Mental Disorders* (5th edition; *DSM-5*; American Psychiatric Association [APA], 2013) has included CU traits as a specifier for the diagnosis of Conduct Disorder (CD) labelled “with Limited Prosocial Emotions” (LPE). To qualify for the specifier, the person must show two or more of the following CU symptoms over at least 12 months across multiple relationships and settings: (a) lack of remorse or guilt; (b) callous-lack of empathy; (c) unconcerned about performance; and (d) shallow or deficient affect (APA, 2013). As noted in the *DSM-5*, for these symptoms to be indicative of the specifier, they must “reflect the individual’s typical pattern of emotional and interpersonal functioning and not just occasional occurrences in some situations” (p. 47, APA, 2013). Further, the World Health

Organization (WHO) added a similar specifier for the diagnoses of both Oppositional Defiant Disorder (ODD) and Conduct-Dissocial Disorder in the latest edition of the International Classification of Disease (11th edition; *ICD-11*; WHO, 2018). In addition to the use of the specifier with the diagnosis of ODD, which is not allowed in the DSM-5, the ICD-11 criteria also includes an additional symptom describing “a relative indifference to the probability of punishment” that is not included in the DSM-5 criteria.

The adoption of this specifier in the two major systems for diagnosing severe behavior problems has led to an increased focus on how to measure these traits in many different clinical and forensic contexts. To date, CU traits have most often been assessed in research using rating scales completed by the person being evaluated or by significant others, such as a child’s parents and teachers (Kotler & McMahon, 2005; Sharp & Kine, 2008). One of the most commonly used measures of CU traits in research is the Inventory of Callous-Unemotional Traits (ICU; Kimonis et al., 2008). The ICU is a 24-item behavior rating scale that includes forms for self-report, as well as parent and teacher ratings. The ICU was developed to (a) provide a focused and comprehensive assessment of CU traits only and not the other dimensions of psychopathy (interpersonal, behavioral, and antisocial facets; Hare, 2003); (b) include a rating format that allows for sufficient variability in responses but does not include a central tendency point (i.e., items are anchored on a 4-point Likert scale from 0 *Not at all true* to 3 *Definitely true*); and (c) include an equal numbers of items that are positively and negatively worded (Frick & Ray, 2015). To date, the ICU has been translated into over 25 different languages and has been used widely in research, with over 200 published studies in samples ranging in age from 3 years to young adulthood, which provide data to support its ability to differentiate both clinically and etiologically important subgroups of children and adolescents with severe behavior problems (Cardinale & Marsh, 2017; Frick

& Ray, 2015). Thus, the ICU has proven to be a time-efficient, reliable, and valid tool for assessing CU traits in a wide range of samples and in many research contexts.

Others behavior rating scales have also been used to assess CU traits in children and adolescents, such as the Youth Psychopathic traits Inventory (YPI; Andershed, Kerr, Stattin, & Levander, 2002) and the Antisocial Process Screening Device (APSD; Frick & Hare, 2001). These rating scales have a less extensive item pool for measuring CU traits than the ICU. Also, the YPI assesses CU traits only via self-report and has no items measuring the symptom “unconcerned about performance” (Andershed et al., 2002). Finally, the limited number of items, combined with the low base rate of endorsement of the items in most settings, has led to low internal consistency for the CU subscales from the YPI and the APSD in many studies (e.g., Poythress, Dembo, Wareham, & Greenbaum, 2006).

While there has been some success in assessing CU traits using behavior rating scales, their usefulness for evaluating the LPE specifier is limited in several ways. First, with the exception of the ICU, most rating scales do not measure CU traits as they are defined by the LPE specifier in the DSM-5. Second, while rating scales typically capture the frequency of CU traits, they often do not allow for the assessment of the persistence (e.g., at least 12 months) and pervasiveness (e.g., typical pattern of functioning across relationships and settings) of the traits that is required by the specifier. Third, although there have been attempts to use behavior ratings to determine “elevations” on CU traits to approximate the LPE specifier, there has not been a single method that has proven valid across informants and samples (Docherty, Boxer, Huessmann, O’Brien, & Bushman, 2016; Kimonis, Fanti, & Singh, 2014; Kimonis et al., 2015). Fourth, making clinical decisions in child and adolescent psychopathology requires integrating information from multiple sources, given that the signs and symptoms of disorders may not be noticeable to all potential informants (De los Reyes et al., 2015). The need for multiple sources is especially important when assessing persons in

forensic samples, where motivation for accurate reporting may be low and motivation for deception and manipulation high (Kelsey, Rogers, & Robinson, 2015). Finally, a clinical interview allows the assessor to determine if the person being assessed understood questions and is answering questions in the way they are intended, again which is particularly important in forensic samples where there may be an overrepresentation of persons with limited verbal abilities (Vermeiren, De Clipelle, Schwab-Stone, Ruchkin, & Deboutte, 2002).

One possible method for assessing these traits in forensic settings is the Psychopathy Checklist: Youth Version (PCL:YV; Forth, Kosson, & Hare, 2003), which is a clinician rating of 20 traits associated with the construct of psychopathy that includes all items that form the LPE specifier except “unconcerned about performance”. However, because it measures the full construct of psychopathy, the PCL-YV includes a large number of items not included in the specifier, making it very time intensive when the only need is measuring the LPE symptoms. Further, evidence for the validity of the PCL:YV is limited in non-institutionalized community samples and in children below the age of 14 (Kotler & McMahon, 2005; Sharp & Kine, 2008).

As a result of these limitations in the available methods for evaluating the LPE specifier, the Clinical Assessment of Prosocial Emotions, Version 1.1 (CAPE 1.1; Frick, 2013) was developed. The CAPE utilizes the structured professional judgment method to assess the four symptoms of the LPE specifier included in the DSM-5, whereby clinicians utilize all available sources of clinical information, which at the minimum include semi-structured interviews with the child and at least one other informant, to rate prototype descriptions for each symptom in the *DSM-5* specifier. These prototypes include descriptions of each of the LPE symptoms, including their persistence and pervasiveness. Based on manualized training, the clinician rates each LPE symptom on a 3-point scale (0 = *Not Descriptive or Mildly Descriptive*; 1 = *Moderately Descriptive*; 2 = *Highly Descriptive*). To

approximate the *DSM-5* criteria, the LPE specifier is present when two or more items are rated as “2”. Thus, this method allows for the assessor to obtain more detailed information than can be gained from rating scales in order to make diagnostic decisions based on the *DSM-5* criteria for the LPE specifier.

This Current Study

Thus, the CAPE 1.1 provides a potential method for assessing the CU traits included in the *DSM-5* LPE specifier that, while more time intensive than rating scales, could be more appropriate for many clinical and forensic samples. However, to date there has not been a published test of the reliability and validity of the CAPE 1.1, although it is currently being tested in multiple countries (Frick, 2013). Thus, we provide one of the first tests of the psychometric properties (reliability and validity) of the CAPE 1.1 in a sample of young males who were incarcerated in two juvenile detention centers in Catalonia, Spain. First, we tested the internal consistency and inter-rater agreement of the CAPE 1.1. Given that the CAPE 1.1 relies substantially on clinical decision-making to make ratings, albeit with very clear guidelines to follow, tests of inter-rater agreement are particularly important. Second, we tested the association of the CAPE 1.1 scores with self-report and informant ratings of CU traits evaluated by means of the ICU and the YPI, as well as clinician ratings on the PCL-YV. These are important tests of the convergent validity of the CAPE 1.1 with some of the most widely used methods for assessing CU traits in past research. Third, we tested whether the ratings from the CAPE 1.1 are more strongly associated with the callous-unemotional dimension of measures of psychopathy than other dimensions of this construct, given the goal of this measure was to capture only the affective components of this broader construct, as evidence of the divergent validity. Finally, we tested the association of the CAPE 1.1 with measures of externalizing and internalizing problems, with predicted positive associations with externalizing problems and with no associations predicted with internalizing problems.

Method

Participants

The participants of this study were recruited from two secure juvenile detention centers of the Department of Penal Execution in the Community and Juvenile Justice (DGEPCJJ) of the *Generalitat* of Catalonia (Spain). The DGEPCJJ has seven juvenile detention centers for adolescents who had committed offenses between the ages 14 to 18 years. Three of them offer detention in a secure facility. However, one of them is the initial referral center from which they are distributed to the other more long-term facilities according to the sanction. As the first facility is a short-term stay center, participants in the current study were housed at the two long term facilities. A total of 72 juvenile males aged 14 to 22 years ($M = 17.74$; $SD = 1.20$) who were incarcerated in two detention centers, the ‘L’Alzina’¹ ($n = 46$; 63.9%) and the ‘El Segre’ ($n = 26$; 36.1%), participated (about 70% of the total census of the two facilities during of the time of the study). Of note, some participants were older than 18 because they were still in custody for charges they incurred as youth. Seventy-five percent of the sample were born out of Spain; 72.2% of both parents were born out of Spain; and 81.9 % had not completed compulsory secondary education. As the participation of families is highly difficult in forensic samples, the collaboration of the juvenile justice professional responsible for the supervision of the participant at the juvenile justice center (social educator) was used as the informant for scoring both the CAPE 1.1 and the ratings used to validate scores from this measure. Therefore, a minimum of three months under the direct care of the social educator prior to the evaluation was required in order to ensure a minimum knowledge of the participant. The exclusion criterion was the presence of a medical condition that contraindicated participation in the study ($n = 1$). Participation was voluntary and anonymous. Participants received a pair of earphones at the L’Alzina center and a computer flash drive at the El Segre center as incentives to participate.²

Materials

Sociodemographic information. Information was obtained through the official records from the juvenile justice center and includes the following variables: age, origin, family origin, and school level reached.

Clinical Assessment of Prosocial Emotions. Version 1.1 (CAPE 1.1; Frick, 2013). The CAPE 1.1 is a clinician rating designed specifically to assess the *DSM-5* LPE specifier for the diagnostic criteria for CD in persons from ages 3 to 21. The four criteria indicated by the *DSM-5* are evaluated: (a) lack of remorse or guilt; (b) callous-lack of empathy; (c) unconcerned about performance; (d) shallow or deficient affect. It utilizes the structured professional judgment method where prototypes for each key indicator of CU traits are provided in order to guide the clinician using the tool. It is designed to be used by clinicians with experience and knowledge of the assessment of childhood psychopathology in general and of CU traits specifically. Clinical judgements are based on multiple informants and sources of information (e.g., records, observation). The CAPE 1.1 includes two semi-structured interviews: Informant Interview and a Self-Report Interview. The interviews include two or three items rating each of the 4 criteria for the LPE specifier. Each item starts with a stem questions (e.g., “Does _____ seem to feel bad or guilty if he/she does something wrong or if he/she hurts someone? “) that must be answered as either “yes” or “no” by the informant. The stem questions are followed by a request for examples during which the clinician can ask whatever follow-up questions she/he feels are indicated to gain enough information to make ratings of the four symptoms. Stem questions are also followed by questions assessing how typical the characteristic might be of the child (e.g., “Is this how he/she is most of the time and with most people?”). The administration of both interviews and the clinician ratings of the four symptoms took between 90 to 120 minutes. The interviews were recorded (audio), except for one participant who did not consent to the recording.

In this study, we used the Catalan and Spanish versions of the CAPE 1.1. The research team translated the symptom descriptions and the two semi-structured interviews into Spanish and the translation was reviewed by other members. A back-translation was made by a specialized translator of the foreign language service of one of the participating institutions. The back-translation was reviewed by the research team (which includes the CAPE's creator) and minor changes were applied. Subsequently, the same procedure was carried out for the Catalan version. Two independent clinicians scored the CAPE 1.1 (interviewer and observer) at each juvenile justice center. All clinicians received individual and group training sessions following the training guidelines provided for the CAPE 1.1 (Frick, 2013). Training sessions were conducted by the first author (BM) who in turn had received training from the CAPE's creator (PJF). First, a one-day workshop was presented by BM to provide clinicians with training and practice in the administration, scoring, and interpretation of the CAPE 1.1. In this session, theoretical material about CU traits and the CAPE 1.1 was provided for individual study and participants discussed a prototype of a person with CU traits from a TV series and scored the CAPE 1.1. Second, interviewers started to administer the CAPE 1.1 at each center. All clinicians scored the CAPE 1.1 of the first two cases from each juvenile justice center (a total of four cases) and provided comments about the administration. Every clinician received individual feedback about their interviews and CAPE 1.1 scoring from BM. Third, clinicians discussed by phone the administration and scoring of the four cases. Fourth, BM supervised the subsequent CAPE's administrations and scorings by reviewing the audio recordings. As required for the CAPE 1.1 (Frick, 2013), clinicians had significant experience in the clinical assessment of psychopathology in children and adolescents and were bilingual in Catalan and Spanish. As both languages are official in Catalonia, the linguistic criteria used was the interviewer's response (social educator/self) regarding the language of preference to communicate more comfortably during the interview.

Inventory of Callous-Unemotional Traits (ICU; Kimonis et al., 2008). The ICU is a 24-item (e.g., “I do not show my emotions to others”) rating scale designed to be a comprehensive measure of CU traits. In this study, we used the authorized Spanish version of the “Unitat d’Epidemiologia i Diagnòstic en Psicopatologia del Desenvolupament” of the Autonomous University of Barcelona (Ezpeleta, de la Osa, Granero, Penelo, & Domènech, 2013). Items are rated on a 4-point Likert scale from 0 (*Not at all true*) to 3 (*Definitely true*). In this study we used the teacher and self-report versions, with the former completed by the participant’s social educators. As recommended by Ray et al. (2016) and consistent with the majority of research using the ICU (Cardinale & Marsh, 2017; Frick & Ray, 2015), the total score of the ICU was used. This method was also justified by the goal of the CAPE 1.1, which is to assess the overall construct of CU traits. In the current sample, the Cronbach’s alpha for the total scale was .77 (mean inter-item correlation [MIC] = .13) for the self-report version and .90 (MIC = .29) for the teacher version completed by the social educator.

Youth Psychopathic traits Inventory (YPI). The YPI (Andershed et al., 2002) is a 50-item self-report questionnaire that measures the core features of psychopathy in adolescents and young adults. It consists of 50 items distributed into 10 subscales and three factors: Grandiose-Manipulative (including dishonest charm, grandiosity, lying and manipulation), Callous-Unemotional (including remorselessness, unemotionality and callousness) and Impulsive-Irresponsible (including thrill-seeking, impulsivity and irresponsibility). Items are rated on a 4-point Likert scale from 1 (*Does not apply at all*) to 4 (*Applies very well*). In the current study, the Cronbach’s alpha for the Total scale was .89 (MIC = .14), for the Grandiose-Manipulative subscale was .90 (MIC = .32), for the Callous-Unemotional subscale was .70 (MIC = .14), and for the Impulsive-Irresponsible subscale was .69 (MIC = .13) for the total sample.

Psychopathy Checklist - Youth Version (PCL-YV). The PCL:YV is a 20-item clinician rating scale that measures psychopathic traits in adolescents and young adults (Forth et al., 2003). It consists of a semi-structured interview and collateral information for assessing interpersonal and affective characteristics and the irresponsible and antisocial lifestyle dimensions of psychopathy (Hare, 2003). Each item is scored by an expert on a 3-point scale ranging from 0 (*Does not apply at all*) to 2 (*Definitely applies*). In the present study, we used the official authorized Spanish translation (Molinuevo, Pardo, González, & Torrubia, 2014). The interview was administered by two clinicians independently at each juvenile justice center and averaged scores were used in statistical analyses. Inter-rater agreement in the present study was calculated at each center (model two-way mixed, type absolute agreement and average measures). In the current research, the Cronbach's alpha for the Total Score was .90 (MIC = .36), for the Interpersonal facet (Facet 1) was .83 (MIC = .54), for the Affective facet (Facet 2) was .82 (MIC = .57), for the Behavioral facet (Facet 3) was .79 (MIC = .43), and for the Antisocial facet (Facet 4) was .72 (MIC = .37).

Achenbach System of Empirically Based Assessment (ASEBA; Achenbach & Rescorla, 2001). The Youth Self Report (YSR) and the Teacher's Report Form (TRF) were used to measure internalizing and externalizing problems. In this study, the TRF was completed by the participant's social educators. They are composed of 112 items that have to be answered on a 3-point scale ranging from 0 (*Not at all*) to 2 (*Often*), with higher scores indicating more problems. To assess internalizing problems, the syndrome subscales of Anxious/Depressed, Withdrawn/Depressed, and Somatic Complaints were used. To assess externalizing problems, the syndrome subscales of Rule-Breaking Behavior and Aggressive Behavior were used. The Spanish adaptation of the instrument has shown satisfactory psychometric properties (Abad, Forns, Amador, & Martorell, 2000). In the current research, the Cronbach's alphas for the Anxious/Depressed subscale were .61 and .72 (MIC = .11 /

.19), for the Withdrawn/Depressed subscale were .77 and .55 ($MIC = .31 / .13$), for the Somatic Complaints subscale were .65 and .77 ($MIC = .21 / .27$), for the Rule-Breaking Behavior subscale were .77 and .78 ($MIC = .18 / .25$), and for the Aggressive Behavior subscale were .78 and .93 ($MIC = .18 / .38$) for the YSR and the TRF forms, respectively.

Procedure

Once the study was approved by the Animal and Human Experimentation Ethics Committee of the Autonomous University of Barcelona and received the authorization of the DGEPCJJ, it was introduced to the directors of the two collaborating juvenile detention centers. After their approval, an information session for the juvenile justice professionals of each center was held. Data collection began in February 2016 and lasted one year and a half. All those adolescents who were incarcerated at each center who met the inclusion criteria were invited to participate. Participant, social educator and parental (for minors) written informed consent were obtained. Though all youth agreed to participate, not all were evaluated due to difficulties in scheduling data collection due to release, trials, or escapes.

The testing of participants was carried out in four sessions, two group (Sessions 1 and 4) and two individual (Sessions 2 and 3) sessions. The YSR was administered in Session 1, the CAPE 1.1 and the ICU in Session 2, the PCL:YV in Session 3, and the YPI in Session 4. Most group sessions were held with 4-5 participants and most sessions were conducted within one month (one session per week). However, some sessions were delayed to accommodate to the needs of the centers. The information from the social educators (CAPE 1.1, ICU, and TRF) was collected individually in one session.

In order to calculate the CAPE 1.1 scores, the CAPE 1.1 semi-structured interviews were separately administered to the participants and to social educators. Two raters independently scored the CAPE 1.1 based on the interviews and file information from the justice juvenile center. Thus, the CAPE 1.1 was scored using the same information at both

centers. The procedure used for the interview at 'L'Alzina' center, involved a rater conducting the interview and the second rater acting as an observer for the 64% of the cases. At 'El Segre' center, only the interviewer was present during the interview. Accordingly, the second rater evaluated 36% of the cases from 'L'Alzina' and 100% of the cases from 'El Segre' by using the audio recordings. The second rater was the same in both centers. As regards to the PCL:YV, two raters, who were not the raters of the CAPE 1.1, independently scored each participant. The procedure involved one rater conducting the interview (recorded by audio), while a second rater was an observer in 55.3 % of cases from 'l'Alzina' and in 100% of cases from 'El Segre'. Revision of file information was made independently by each rater. Accordingly, 44.7% of the cases from 'l'Alzina' were evaluated by the second rater by using the audio recording and collateral information. All the raters had been previously trained in the use of the CAPE 1.1 and the PCL:YV and either held a PhD, had clinical experience, or were advanced doctoral students in psychology.

Statistical analyses

First, we tested the reliability of the CAPE 1.1 in several ways. Cohen's kappa (κ) was calculated to test interrater agreement of the CAPE 1.1 by measuring the concordance between interviewer and observer on the LPE diagnosis from the CAPE 1.1. Guidelines provided by Landis and Koch (1977) are as follows: values from 0.00 to 0.20 indicate slight agreement, 0.21 to 0.40 fair agreement, 0.41 to 0.60 moderate agreement, 0.61–0.80 substantial agreement, and 0.81–1.00 almost perfect agreement. To evaluate the internal consistency of the CAPE 1.1, the Cronbach's alphas were calculated and interpreted as poor ($\leq .60$), marginal (.60 to .69), acceptable (.70 to .79), good (.80 to .89), and excellent ($\geq .90$; Barker, Pristang, & Elliott, 2002). Additionally, because of the dependence of Cronbach's α on the number of items of a scale and the very few items on the CAPE 1.1 (i.e., four), MIC was also provided as an indicator of the internal consistency, with values ranging .15 to .50

being considered adequate (Clark & Watson, 1995). Second, we estimated the prevalence of each of the four symptoms of the specifier rated “2” (*Highly Descriptive*) and the prevalence of youth meeting the criteria for the specifier (2 or more symptoms rated as “2”) in the full sample according to the interviewer. Third, to test the convergent and the divergent validity of the CAPE 1.1 we performed tests for independent samples comparing groups with and without LPE on other measures of CU traits (ICU) and psychopathy (YPI and PCL:YV) and of externalizing and internalizing problems (TRF/YSR). Because the groups are not large, Shapiro-Wilk tests of normality were performed. In case of normality, Student’s *t*-test were computed and the effect size of mean comparison was estimated using Cohen’s *d*. This index was judged as small: $d = .2$, medium: $d = .5$, and large: $d = .8$ (Cohen, 1992). In case of non-normality, Mann-Whitney U test were performed and the effect size of rank-sum comparison was estimated using *r*. This index was judged as small: $r = .1$, medium: $r = .3$, large: $r = .5$, and very large: $r = .7$ (Cohen, 1992). Statistical analyses for each rating scale were performed by only including cases that had answered at least 70% of the items. All analyses were conducted using IBM SPSS Statistics Version 21 (SPSS, Inc., Chicago, IL, USA).

Results

Reliability

Table 1 presents the inter-rater agreement between interviewers and observers for each CAPE 1.1 item and for the LPE specifier. The interrater agreement on the cases meeting the LPE criteria (i.e., 2 or more symptoms rated as ‘Highly Descriptive’) was .66 indicating substantial agreement for this score, which is the primary variable of interest from the CAPE 1.1. The inter-rater agreement of the individual symptoms ranged from .30 to .51 indicating moderate agreement for the items 3 ‘Unconcerned about Performance’ and 2 ‘Lack of Empathy’, and fair for the items 4 ‘Shallow or Deficient Affect’ and 1 ‘Lack of Remorse or

Guilt'. The Cronbach's alpha of the CAPE 1.1 was .81 and the MIC was .51. These indices suggest a good level of internal consistency across symptoms.

Prevalence

Table 2 shows the prevalence of each CAPE 1.1 symptom rated as "Highly Descriptive" and the prevalence of the LPE specifier for the full sample. As shown, the LPE specifier was present in more than half of the sample (55.6%) and the two items most commonly endorsed as being "Highly Descriptive" were item 1 'Lack of Remorse or Guilt' (59.7%) and item 2 'Lack of Empathy' (54.2%). The prevalence of the number of items scored as "Highly Descriptive" were as follow: 31.9% of youth ($n = 23$) had zero items, 12.5% ($n = 9$) one item, 18.1% ($n = 13$) two items, 26.4% ($n = 19$) three items, and 11.1% ($n = 8$) four items.

Convergent and Divergent Validity

Youth who met LPE diagnostic criteria assessed by the CAPE 1.1 scored significantly higher on CU traits measured by the ICU compared to those not meeting criteria (Table 3). The two groups differed on the total scores for the ICU according to both informants: the social educator ($d = 1.01$) and the youth ($d = 0.57$). These scores indicate that there was relatively strong convergence between the CAPE 1.1 and this rating scale measure of CU traits, especially according to the social educator scores.

Youth who met LPE diagnostic criteria assessed by the CAPE 1.1 did not score significantly higher on CU traits as measured by the YPI. However, as indicated in Table 3, youth who met LPE diagnostic criteria assessed by the CAPE 1.1 scored significantly higher on CU traits measured by the PCL:YV (Facet 2 - Affective) compared to those not meeting criteria ($r = 0.30$). Significant differences were also found on Facet 1 - Interpersonal ($r = 0.27$) and Facet 4 – Antisocial ($r = 0.25$), but as expected, no significant differences were found on Facet 3 - Behavioral.

Therefore, there was some evidence of convergent (Facet 2 - Affective) and divergent validity (Facet 3 - Behavioral) on the PCL-YV. However, the results showing that those with LPE specifier also scored higher on Facets 1 (Interpersonal) and 4 (Antisocial) were not predicted. However, this finding may have been due to the correlations among these three facets (correlations range = .43 - .58; $p < .01$). Thus, logistic regression analyses were performed considering the specifier LPE measured by the CAPE 1.1 as the outcome variable, and the four PCL:YV facets were entered as independent variables. The forward and backward selection methods (likelihood ratio in both) were computed. Results showed that the Facet 2 (Affective) was the only unique predictor of the LPE specifier (Odds ratio = 1.51; $p = .01$).

Finally, Table 4 shows the comparisons of the participants who met LPE diagnostic criteria to those who did not meet this threshold on the external variables of externalizing and internalizing problems according to the two informants (social educator and youth). Youth who met the LPE measured by the CAPE 1.1 were scored higher on the Aggressive Behavior subscale and on the Rule-Breaking subscale according to both informants. The effect sizes of these differences were moderate (d/r range = .33 - .68). As predicted, no significant differences were found between those meeting and those not meeting the LPE criteria for the internalizing syndrome subscales across both informants. These results support the convergent and divergent validity of the CAPE 1.1.

Discussion

The aim of this study was to provide an initial test of the psychometric properties (reliability and validity) of the Spanish/Catalan version of the CAPE 1.1 in a sample of young males who were incarcerated. This is an important initial test of the only clinician rating that has been developed specifically to measure the symptoms of the new LPE specifier, which

was recently included as a part of the diagnostic criteria for CD in the *DSM-5* (APA, 2013). While the CU traits that make up this specifier have been widely measured in research using rating scales, their inclusion in the diagnostic criteria for the first time requires the development and testing of more comprehensive methods of assessment that are appropriate for making decisions on clinical diagnoses (Frick & Ray, 2015).

In terms of the reliability of the CAPE 1.1, the diagnosis of LPE showed adequate interrater agreement and internal consistency. This is important because this is the primary clinical decision for which the CAPE 1.1 was developed. Further, these results support past work suggesting that structured clinical decisions that weigh information from multiple sources can be made reliably when clinicians are appropriately trained and clear rating guidelines are provided (Hilterman, Nicholls, & van Nieuwenhuizen, 2013). Of note, the individual symptoms showed much more modest levels of interrater agreement, calling into question interpretations made from these individual symptoms. The fact that the second rater (i.e., observer) scored more than a half interviews listening the audios and consulting collateral information could partly explain these modest levels of reliability. It is possible that direct observation of the person's behavior, especially the nonverbal behavior, may be a relevant source of information for the clinician, especially when rating specific affective traits (Knapp & Hall, 2010; Kosson, Gacono, Klipfel, & Bodholdt, 2016). While the actual reason for the lower agreement in some particular CAPE 1.1 symptoms cannot be conclusively determined in the current study, it does suggest that future research should test potential influences on the consistency of clinical ratings, including the reliability and validity of scoring the CAPE 1.1 based on audio or video recordings.

The limited research on the prevalence of the new *DSM-5* LPE specifier, together with variations in how it has been measured in past studies, makes it difficult to determine what would be an appropriate prevalence rate with which to compare the rate of 56.5% found in

the current sample of detained adolescent boys. When considering studies using rating scales to assess the specifier, they often did not measure the symptom “unconcerned about performance” and/or did not refer to any time frame in particular, like the 12 months specified by the *DSM-5* temporal criterion (Colins, 2016; Colins & Vermeiren, 2013). Also, prevalence rates from rating scales differ according to the characteristics of the sample (e.g., forensic, clinical, community), the informant (self-report vs. others), and the criteria for considering a symptom to be present based on the ratings (see Kimonis et al., 2015 for a discussion of this last fact). As would be expected, the rates of children and adolescents meeting the threshold for the specifier tends to be much lower in community samples, with rates typically being below 5% (Kahn, Frick, Youngstrom, Findling, & Youngstrom, 2012; McMahon, Witkiewitz, Kotler, & the Conduct Problems Prevention Research Group, 2010; Pardini, Stepp, Hipwell, Stouthamer-Loeber, & Loeber, 2012; Seijas, Servera, García-Banda, Barry, & Burns, 2018). The rates are higher in forensic samples but depend on the informant. For example, Van Damme, Colins, & Vanderplasschen (2016) reported higher rates in girls who were incarcerated when considering parents’ report (56.5%) vs. youth report (17.6%). Kimonis et al. (2014) reported prevalence rates of 15.8% when relying on self-report in adolescent boys who were incarcerated. Vanwoerden, Reuter, & Sharp (2016) reported that 16.2% of adolescents on a psychiatric inpatient unit met the specifier requiring the top rating to be considered indicative of the symptom but a prevalence rate of 67.9% using a less stringent method. In summary, more research is needed to test the influences on the various methods of assessing the prevalence of the specifier, although these early findings suggest that reliance only on one informant, notably youth, may significantly underestimate its prevalence. It is possible that methods, such as the CAPE 1.1, which does not rely on any single informant, would result in more consistent prevalence rates.

The results of our study support the convergent validity of the CAPE 1.1 for measuring the *DSM-5* LPE specifier. First, as expected, participants meeting criteria for the LPE specifier compared to those who do not reach the threshold showed higher scores on CU traits according to several raters, informants, and methodologies (interview vs. self-report). The CAPE 1.1 converge with the ICU and the Facet 2 - Affective of the PCL:YV, with the former being based on self and social educator report and the latter being based on clinician ratings. This finding is very important because it links the CAPE 1.1 scores to methods that have been widely used in previous research, which has documented the association between CU traits with clinically important outcomes in children, adolescents, and young adults (Frick et al., 2014b).

With regards to divergent validity, we showed that diagnoses on the CAPE 1.1 were more consistently related to measures of CU traits relative to other dimensions of psychopathy. While those diagnosed with the LPE specifier still differed on the Facet 1 and Facet 4 of the PCL-YV, our results showed that this could reflect correlations among the dimensions that form the construct of psychopathy, given that this association was no longer present when controlling for the other dimensions of psychopathy (Drislane & Patrick, 2017). In addition, as expected, results suggest that the CAPE 1.1 detects a subgroup of youth with more severe externalizing problems such as rule-breaking and aggressive behavior but who do not present with high levels of internalizing problems (Cardinale & Marsh, 2017).

Importantly, future research needs to determine if the CAPE 1.1 diagnoses add to the prediction of important outcomes relative to existing rating scales. That is, the justification for the development of the CAPE 1.1 is that it allows for a structured method for weighing multiple sources of information to make an important clinical decision. However, it would be essential to determine if this more time-consuming method provides incremental utility to

more time efficient rating scales for predicting important clinical outcomes (e.g., impairment, risk for later antisocial behavior, and response to treatment).

Limitations

The current study needs to be considered in the context of certain limitations. First, and most importantly, it did not include an assessment of the diagnosis of CD or ODD, which is required by the *DSM-5* and/or the *ICD-11* diagnostic criteria. However, all youth included in the present study were adjudicated for illegal behavior (e.g., robbery with/without violence, drug offenses, assault, sexual offenses, or murder) and, for more than a third of them, it was not the first time that they were detained . Thus, it is likely that a majority of participants would have met criteria for CD or ODD at some point in their lives. Nevertheless, future research should include a formal method for evaluating the diagnosis of CD and ODD, to correspond more directly with how the LPE is used in the *DSM-5* and *ICD-11*³. Second, the use of a sample of males from Spain who were incarcerated means that these results may not be generalizable to females; to children, adolescents and young adults from community and clinical samples; samples from other cultural backgrounds; and persons in other juvenile justice contexts that do not require detention. Third, and also related to the use of a detained sample, the scoring of the CAPE 1.1 relied on the report of the participant and his teacher (social educator). While this follows the requirements of the CAPE 1.1 to gain information from multiple informants and the participant's teacher had substantial contact with the participant in the detention facility, the absence of information from the participant's parents may have decreased the validity of the clinical diagnoses. Finally, the small sample size of the study is another limitation and future research is needed to further evaluate the psychometric properties of the CAPE 1.1 in larger samples.

Conclusions and Implications

In the context of these limitations, our data provide some initial and promising data to support the psychometric properties of the CAPE 1.1. Most importantly, these results suggest that the CAPE 1.1, which relies on trained clinicians to make structure judgements, leads to reliable diagnoses of the *DSM-5* LPE specifier. As noted above, this specifier leads to both clinically and etiologically important subgroups of children and adolescents with serious behavior problems. Further, our findings suggest that the diagnoses that result from the CAPE 1.1 are highly related to measures that have been used to rate CU traits in past research. Clearly more work is needed to establish what interpretations can be validly made from the CAPE 1.1 and whether this time-consuming clinical procedure provides important information that cannot be obtained in more economical formats. However, these results provide an important first step in this process of developing a tool for assessing a construct that is relatively new to mental health diagnoses of children and adolescents in a way that may be useful in many clinical and forensic settings in culturally diverse samples.

References

- Abad, J., Forns, M., Amador, J. A., & Martorell, B. (2000). Fiabilidad y validez del youth self report en una muestra de adolescentes [Youth Self-Report reliability and validity in an adolescent sample]. *Psicothema*, 12(1), 49-54.
- Achenbach, T. M., & Rescorla, L. A. (2001). Manual for the ASEBA School-Age forms and profiles. Burlington, VT: University of Vermont Research Center for Children, Youth, & Families.
- Andershed, H., Kerr, M., Stattin, H., & Levander, S. (2002). Psychopathic traits in non-referred youths: Initial test of a new assessment tool. In E. Blaauw & L. Sheridan (Eds.), *Psychopaths: Current international perspectives* (pp. 131–158). The Hague: Elsevier.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Blair, R. J. R., Leibenluft, E., & Pine, D. S. (2014). Conduct disorder and callous–unemotional traits in youth. *New England Journal of Medicine*, 371(23), 2207–2216. doi:10.1056/NEJMra1315612
- Barker, C., Pristang, N., & Elliott, R. (2002). *Research methods in clinical psychology: An introduction for students and practitioners. The marketing review* (2nd ed.). Chichester, UK: Wiley.
- Cardinale, E. M., & Marsh, A. A. (2017). The reliability and validity of the Inventory of Callous Unemotional Traits: a meta-analytic review. *Assessment*. Advance online publication. doi:1073191117747392
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319. doi:10.1037/1040-3590.7.3.309

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.

doi:10.1037/0033-2909.112.1.155

Colins, O. F. (2016). The clinical usefulness of the DSM-5 specifier for conduct disorder outside of a research context. *Law and Human Behavior*, 40(3), 310-318.

doi:10.1037/lhb0000173

Colins, O. F., & Vermeiren, R. R. J. (2013). The usefulness of DSM-IV and DSM-5 conduct disorder subtyping in detained adolescents. *The Journal of Nervous and Mental Disease*, 201(9), 736-743. doi:10.1097/NMD.0b013e3182a20e94

De los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, 141(4), 858-900.

doi:10.1037/a0038498

Docherty, M., Boxer, P., Huesmann, L. R., O'Brien, M., & Bushman, B. (2017). Assessing callous-unemotional traits in adolescents: Determining cutoff scores for the Inventory of Callous and Unemotional traits. *Journal of Clinical Psychology*, 73(3), 257-278.

doi:10.1002/jclp.22313

Drislane, L. E., & Patrick, C. J. (2017). Integrating alternative conceptions of psychopathic personality: A latent variable model of triarchic psychopathy constructs. *Journal of Personality Disorders*, 31(1), 110–132. doi:10.1521/pedi_2016_30_240

Ezpeleta, L., de la Osa, N., Granero, R., Penelo, E., & Domènech, J. M. (2013). Inventory of Callous-unemotional traits in a community sample of preschoolers. *Journal of Clinical Child and Adolescent Psychology*, 42(1), 91–105. doi:10.1080/15374416.2012.734221

Forth, A., Kosson, D., & Hare, R. D. (2003). *Psychopathy Checklist: Youth Version*. Toronto, Canada: Multi-Health Systems.

Frick, P. J. (2013). *Clinical Assessment of Prosocial Emotions: Version 1.1 (CAPE 1.1)*.

[Measurement instrument]. Retrieved from <https://sites01.lsu.edu/faculty/pfricklab/wp-content/uploads/sites/100/2015/11/CAPE-Manual.pdf>

Frick, P., & Hare, R. D. (2001). *The Antisocial Processes Screening Device: Technical Manual*. Toronto, ON: Multi-Health Systems.

Frick, P. J., & Ray, J. V. (2015). Evaluating Callous-Unemotional Traits as a Personality Construct. *Journal of Personality*, 83(6), 710–722. doi:10.1111/jopy.12114

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014a). Annual research review: A developmental psychopathology approach to understanding callous-unemotional traits in children and adolescents with serious conduct problems. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 55(6), 532–548. doi:10.1111/jcpp.12152

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014b). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin*, 140(1), 1–57. doi:10.1037/a0033076

Frick, P. J., Stickle, T. R., Dandreaux, D. M., Farrell, J. M., & Kimonis, E. R. (2005). Callous-unemotional traits in predicting the severity and stability of conduct problems and delinquency. *Journal of Abnormal Child Psychology*, 33(4), 471–487. doi:10.1007/s10648-005-5728-9

Hare, R. D. (2003). *The Hare Psychopathy Checklist-Technical manual (2nd ed.)*. Toronto: Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a clinical and empirical construct. *Annual Review of Clinical Psychology*, 4, 217–246. doi:10.1146/annurev.clinpsy.3.022806.091452

- Herpers, P. C. M., Rommelse, N. N. J., Bons, D. M. A., Buitelaar, J. K., & Scheepers, F. E. (2012). Callous-unemotional traits as a cross-disorders construct. *Social Psychiatry and Psychiatric Epidemiology*, 47(12), 2045–64. doi:10.1007/s00127-012-0513-x
- Hilterman, E. L. B., Nicholls, T. L., & van Nieuwenhuizen, C. (2013). Predictive validity of risk assessments in juvenile offenders: Comparing the SAVRY, PCL:YV, and YLS/CMI with unstructured clinical assessments. *Assessment*, 21(3), 324–339. doi:10.1177/1073191113498113
- Kahn, R. E., Frick, P. J., Youngstrom, E., Findling, R. L., & Youngstrom, J. K. (2012). The effects of including a callous-unemotional specifier for the diagnosis of conduct disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 53, 271–282. doi:10.1111/j.1469-7610.2011.02463.x
- Kelsey, K. R., Rogers, R., & Robinson, E. V. (2015). Self-Report measures of psychopathy: What is their role in forensic assessments? *Journal of Psychopathology and Behavioral Assessment*, 37(3), 380–391. doi:10.1007/s10862-014-9475-5
- Kimonis, E. R., Fanti, K. A., Frick, P. J., Moffitt, T. E., Essau, C., Bijttebier, P., & Marsee, M. A. (2015). Using self-reported callous-unemotional traits to cross-nationally assess the DSM-5 “With Limited Prosocial Emotions” specifier. *Journal of Child Psychology and Psychiatry*, 56, 1249-1261. doi:10.1111/jcpp.12357
- Kimonis, E. R., Fanti, K., Goldweber, A., Marsee, M. A., Frick, P. J., & Cauffman, E. (2014). Callous-unemotional traits in incarcerated adolescents. *Psychological Assessment*, 26(1), 227–37. doi:10.1037/a0034585
- Kimonis, E. R., Fanti, K. A., & Singh, J. P. (2014). Establishing cut-off scores for the parent-reported Inventory of Callous-Unemotional Traits. *Archives of Forensic Psychology*, 1(1), 27-48.

- Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Munoz, L. C., ... & Morris, A. S. (2008). Assessing callous–unemotional traits in adolescent offenders: Validation of the Inventory of Callous–Unemotional Traits. *International Journal of Law and Psychiatry*, 31, 241–252. doi:10.1016/j.ijlp.2008.04.002
- Knapp, M. L., & Hall, J. A. (2010). *Nonverbal communication in human interaction* (7th ed.). Boston, MA: Wadsworth, Cengage Learning.
- Kosson, D. S., Gacono, C., Klipfel, K. M., & Bodholdt, R. (2016). Understanding and assessing psychopathy: Interpersonal aspects and clinical interviewing. In C. B. Gacono (Ed.), *The clinical and forensic assessment of psychopathy* (2nd ed., pp. 252–275). New York, NY: Routledge, Taylor & Francis Group.
- Kotler, J. S., & McMahon, R. J. (2005). Child psychopathy: Theories, measurement, and relations with the development and persistence of conduct problems. *Clinical Child and Family Psychology Review*, 8(4), 291–325. doi:10.1007/s10567-005-8810-5
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. doi:10.2307/2529310
- Molinuevo, B., Pardo, Y., González, L., & Torrubia, R. (2014). Memories of parenting practices are associated with psychopathy in juvenile male offenders. *The Journal of Forensic Psychiatry & Psychology*, 25(4), 495–500. doi:10.1080/14789949.2014.920901
- McMahon, R. J., Witkiewitz, K., Kotler, J. S., & the Conduct Problems Prevention Research Group. (2010). Predictive validity of callous unemotional traits measured in early adolescence with respect to multiple antisocial outcomes. *Journal of Abnormal Psychology*, 119, 752–763. doi:10.1037/a0020796

- Pardini, D., Stepp, S., Hipwell, A., Stouthamer-Loeber, M., & Loeber, R. (2012). The clinical utility of the proposed DSM-5 callous-unemotional subtype of conduct disorder in young girls. *Journal of the American Academy of Child and Adolescent Psychiatry*, 51(1), 62–73.e4. doi:10.1016/j.jaac.2011.10.005
- Poythress, N., Dembo, R., Wareham, J., & Greenbaum, P. E. (2006). Construct validity of the Youth Psychopathic Traits Inventory (YPI) and the Antisocial Process Screening Device (APSD) with justice-involved adolescents. *Criminal Justice and Behavior*, 33(1), 26–55. doi:10.1177/0093854805282518
- Ray, J. V, Frick, P. J., Thornton, L. C., Stein berg, L., Cauffman, E., Ray, J. V, ... Cauffman, E. (2016). Positive and negative item wording and its influence on the assessment of callous-unemotional traits. *Psychological Assessment*, 28(4), 394–404. doi:10.1037/pas0000183
- Seijas, R., Servera, M., García-Banda, G., Barry, C. T., & Burns, G. L. (2018). Evaluation of a four-item DSM–5 limited prosocial emotions specifier scale within and across settings with Spanish children. *Psychological Assessment*, 30(4), 474–485. doi:10.1037/pas0000496
- Sharp, C., & Kine, S. (2008). The assessment of juvenile psychopathy: Strengths and weaknesses of currently used questionnaire measures. *Child and Adolescent Mental Health*, 13(2), 85–95. doi:10.1111/j.1475-3588.2008.00483.x
- Van Damme, L., Colins, O. F., & Vanderplasschen, W. (2016). The limited prosocial emotions specifier for conduct disorder among detained girls: A multi-informant approach. *Criminal Justice and Behavior*, 43(6), 778–792. doi:10.1177/0093854815613104
- Vanwoerden, S., Reuter, T., & Sharp, C. (2016). Exploring the clinical utility of the DSM-5

conduct disorder specifier of “with limited prosocial emotions” in an adolescent inpatient sample. *Comprehensive Psychiatry*, 69, 116–131.

doi:10.1016/j.comppsy.2016.05.012

Vermeiren, R., De Clippele, A., Schwab-Stone, M., Ruchkin, V., & Deboutte, D. (2002).

Neuropsychological characteristics of three subgroups of Flemish delinquent adolescents. *Neuropsychology*, 16, 49–55. doi:10.1037//0894-4105.16.1.49

World Health Organization. (2018). *International statistical classification of diseases and related health problems* (11th Revision). Retrieved from <https://icd.who.int/browse11/l-m/en>

Endnotes

¹The difference in the number of participants from each center is largely due to their accommodation capacity.

²The choice of incentives was based on the preference of the director of the center.

³We made additional analyses by limiting our sample to those scoring equal to or greater than T-scores of 60 on the Rule-breaking behavior or Aggressive behavior subscales of the YSR version ($n = 47$; 65.3% of the total sample). The results showed that participants who met the LPE assessed by the CAPE 1.1 ($n = 30$) scored higher on (a) the ICU (both informants), (b) the CU scale of the YPI, (c) the Facet 2 and the Total score of the PCL:YV, and (d) the Rule-breaking behavior scale of the TRF version. While these results are promising, it is important to note that these results should not be considered as testing the actual DSM-5 criteria for CD, given the limited coverage of CD symptoms on the ASEBA subscales, the limited time frame assessed (6 months) by the YSR, and concerns over the use of the ASEBA norms for a Spanish sample (i.e., TRF version).

Table 1

Inter-rater agreement on the CAPE 1.1 (Cohen's Kappa) for each item and for the LPE specifier

	Cohen's Kappa
CAPE 1.1	(<i>N</i> = 70)
Item 1. Lack of Remorse or Guilt	.30
Item 2. Lack of Empathy	.51
Item 3. Unconcerned about Performance	.43
Item 4. Shallow or Deficient affect	.40
Limited Prosocial Emotions specifier	.66

Note. CAPE 1.1 = Clinical Assessment of Prosocial Emotions: Version 1.1.

Table 2

Frequencies and percentages of each CAPE 1.1 symptom rated as “Highly Descriptive” and of the specifier (interviewer and full sample)

CAPE 1.1	Highly Descriptive (2)
	<i>n</i> (%)
Items	
Lack of remorse or guilt	43 (59.7)
Lack of empathy	39 (54.2)
Unconcerned about performance	15 (20.8)
Deficient affect	27 (37.5)
LPE specifier (presence)	40 (55.6)

Note. CAPE 1.1 = Clinical Assessment of Prosocial Emotions: Version 1.1.; LPE = Limited Prosocial Emotions.

Table 3

Differences in psychopathic traits according to the presence of the LPE specifier

	CAPE 1.1 (interviewer)		<i>t/z</i>	<i>d/r</i>
	LPE Yes	LPE No		
Psychopathic traits	<i>M (SD)</i>	<i>M (SD)</i>		
	<i>n</i> = 39	<i>n</i> = 31		
ICU Total (Social Educator)	1.62 (.51)	1.16 (.39)	4.15***	1.01
	<i>n</i> = 40	<i>n</i> = 32		
ICU Total (Youth)	1.16 (.37)	0.95 (.37)	2.44*	0.57
	<i>n</i> = 39	<i>n</i> = 31		
YPI Total (Youth)	2.25 (.42)	2.21 (.37)	0.39	
YPI Grandiose	1.96 (.58)	1.99 (.58)	-0.23	
YPI Callous	2.13 (.43)	2.04 (.50)	0.86	
YPI Impulsive	2.76 (.49)	2.69 (.36)	0.68	
	<i>n</i> = 33	<i>n</i> = 24		
PCL:YV Total	28.11 (6.96)	23.00 (7.96)	2.63*	0.69
PCL:YV Interpersonal Facet ^a	5.15 (2.00)	3.98 (2.12)	2.08*	0.27
PCL:YV Affective Facet ^a	6.47 (1.57)	5.06 (2.18)	2.35**	0.30
PCL:YV Behavioral Facet	7.21 (2.08)	6.27 (2.48)	1.58	
PCL:YV Antisocial Facet ^a	6.90 (2.24)	5.83 (2.09)	1.96*	0.25

Note. CAPE 1.1 = Clinical Assessment of Prosocial Emotions: Version 1.1.; ICU = Inventory of Callous-Unemotional traits; YPI = Youth Psychopathy Inventory; PCL:YV = Psychopathy Checklist: Youth Version (PCL:YV).

^aNon-parametric test of group differences (*z*) and effect size (*r*).

p* < .05; *p* < .01; ****p* < .001.

Table 4

Differences in externalizing and internalizing problems according to the presence of the LPE specifier (total sample)

	CAPE 1.1 (interviewer)		<i>t/z</i>	<i>d/r</i>
	LPE Yes	LPE No		
ASEBA	<i>M (SD)</i>	<i>M (SD)</i>		
Teacher Report Form (TRF)	<i>n</i> = 40	<i>n</i> = 29		
Rule-breaking behavior	8.90 (4.63)	6.48 (3.79)	2.30*	0.57
Aggressive behavior ^a	12.03 (9.44)	6.14 (5.35)	2.78**	0.33
Anxious/Depressed	6.40 (4.00)	5.48 (3.73)	0.97	
Withdrawn/Depressed ^a	3.83 (3.30)	3.93 (2.87)	-0.35	
Somatic Complaints ^a	0.70 (1.23)	1.00 (1.96)	- 0.36	
Youth Self-Report (YSR)	<i>n</i> = 38	<i>n</i> = 29		
Rule-breaking behavior	14.32 (5.22)	11.59 (4.31)	2.28*	0.57
Aggressive behavior	13.03 (5.96)	9.34 (4.87)	2.71**	0.68
Anxious/Depressed	7.98 (4.40)	6.97 (2.28)	1.04	
Withdrawn/Depressed ^a	6.11 (3.04)	5.21 (1.54)	1.09	
Somatic Complaints ^a	3.89 (3.78)	3.00 (2.07)	0.43	

Note. CAPE 1.1 = Clinical Assessment of Prosocial Emotions: Version 1.1.; LPE = Limited Prosocial Emotions; ASEBA = Achenbach System of Empirically Based Assessment.

^aNon-parametric test of group differences (*z*) and effect size (*r*).

p* < .05; *p* < .01.